Learning about the Enforcement of Conditional Welfare Programs and Behavioral Responses: Evidence from Bolsa Familia in Brazil

Katja Maria Kaufmann*
Bocconi University and IGIER

Eliana La Ferrara
Bocconi University and IGIER

Fernanda Brollo
University of Alicante

October 2011

PRELIMINARY, PLEASE DO NOT CIRCULATE OR CITE

Abstract

The effectiveness of conditional welfare programs crucially depends on their design, such as the exact rules, benefit amounts and structure. Another important factor is the quality of enforcement of those rules. While the former relationship has been studied, the importance of the quality of enforcement for program effectiveness has rarely been addressed.

In this paper we study the implementation of a large-scale conditional cash transfer program "Bolsa Familia" in Brazil, which conditions transfers to poor families on children's school attendance. We analyze how people learn about the quality of enforcement and how this affects their behavior. We find that individuals respond to incentives and finetune their behavior in response to signals about the quality of enforcement of program conditions. They learn both from private signals, observing the consequences of own noncompliance, and from public signals, that is observing the consequences from peers' noncompliance. Thus we show that enforcement does not only have a direct effect on the family affected, but also an important multiplier effect on other families, who learn from the experiences of their children's peers.

*Corresponding author: Katja Kaufmann, Bocconi University, Via Roentgen 1, 20136 Milan, Italy, E-mail: katja.kaufmann@unibocconi.it

1 Introduction

The effectiveness of welfare programs crucially depends on their design, i.e. on the exact rules, benefit amounts and structure. Another important factor is the quality of enforcement of those rules. While the former relationship has been studied to some extent, for example in the context of conditional cash transfer (CCT) programs (see, e.g. Attanasio, Meghir, and Santiago (2011), Barrera-Osorio, Bertrand, Linden, and Perez-Calle (2011) and Todd and Wolpin (2006)), the importance of the quality of enforcement for program effectiveness has rarely been addressed.

Most of the literature studying the impacts of welfare programs assumes that rules are enforced as specified. In so far as the goal is not only to evaluate a given program for a given level of enforcement, but to offer more general policy recommendations, it is crucial to explicitly take into account the degree of enforcement and to shed light on how the effectiveness of programs varies with the degree of enforcement. Building a reputation of strict enforcement of program rules is an important concern for many developing countries, where welfare programs might get started with insufficient administrative capacity to monitor and implement program conditions.

In this paper we study the structure of conditionality and the quality of enforcement of a large-scale conditional cash transfer program, the Bolsa Familia Program (BFP) in Brazil, which conditions transfers to poor families on children's school attendance. The conditionality scheme of BFP has several interesting features that distinguish it from other CCT programs, including the well-known Progresa/Oportunidades. The first is the joint responsibility of all children in the beneficiary family. In the case of Progresa if a child does not satisfy the schooling requirements in a given month, the family's cash transfer is reduced by the amount that pertains to that child in that month. In the case of Bolsa Familia, on the other hand, *all* children aged 6 to 15 in the family have to attend school a sufficient number of days each month. Failure of any one of the children to comply leads to the loss of benefits for the entire family. This feature of the program is interesting because not only does it constitute a particularly strong incentive to attend school, but it introduces an interdependence among children's attendance within the family that we can exploit.

A second peculiarity of BFP is that its conditionality involves five different stages of warnings with increasing severity of punishment. As we explain in detail in section 2, some of these stages imply a direct loss of money while others don't. This allows us to study how people respond to different warning stages and whether warnings that do not imply a loss of benefit are as effective as warnings that do. This question is of high policy relevance as CCTs typically target families that are severely cash constrained, so evidence that "threats" have an independent effect from financial penalties may be used to design schemes that minimize the adverse impacts of conditionalities on poverty.

We use administrative data on BFP from the Brazilian Ministry of Social Development (MDS) and Ministry of Education (MEC). We have information about child and family characteristics, as well as monthly data on school attendance of each child, monthly information on warnings –from which we create the complete warning history of each family– and monthly information on the benefit the family received (or should have received) and whether the benefit was suspended in a given month. To study how families learn from others about the enforcement of conditionality, we merge the administrative data described above with the School Census data, so that we can identify the BFP recipients' peers who are in the same school, grade and class for each child in the family.

The research questions and findings of this paper can be summarized as follows. First, we ask whether individuals respond differently to different stages of conditionality that they themselves reach. We find that attendance increases in response to warnings received and the more so the higher the stage reached. On average, warnings that involve a financial loss have an impact that is about twice as large as those that do not. The effect is particularly pronounced for liquidity constrained households and increases with the amount of benefit that is potentially lost in case of non-compliance.

Second, we ask whether imperfect enforcement of program conditions weakens the effect on attendance, for a given design of the conditionality structure. We do so in two ways: (i) we test whether warnings have a different effect when they are received with a long delay compared to the month of non-attendance, and (ii) we test whether ex post "justifications" by the school principal or teachers that allow the family to receive the benefit despite noncompliance lead to lower attendance. In both cases we find that the answer is yes, which suggests that families do take into account the quality of enforcement when responding to conditionalities.

Third, we ask whether families learn from the experience of others. We start by testing whether, controlling for a child's own warning stage, warnings received by his/her peers (BFP recipients in the same school, grade or class) increase his/her attendance. We find that they do, and we argue that this is not driven by correlated shocks because of the time lag involved and because persistent shocks would lead to the opposite effect (see section 4 for a detailed discussion of our identification strategy). We also exploit the fact that a child's siblings are typically in a different grade, and sometimes in a different school (depending on age difference), to test whether warnings receives by a sibling's peers also affect a child's attendance. These signals are unlikely to be correlated with potential shocks that may be specific to a child's class (e.g., teacher quality or effort). Interestingly, we find that signals received by siblings' peers do matter, in a qualified sense. When the fraction of the child's own peers that receive the warning is higher than the fraction of siblings' peers, a child's attendance responds more strongly to own peers' signals, as we would expect because that is the strongest signal of enforcement that the child receives. On the other hand, when the fraction of siblings' peers warned is the higher one, then a child's attendance responds more to this signal.

Fourth, as we did for the family's own warnings, we ask whether weak enforcement of conditionalities for a child's peers or his/her siblings' peers leads to lower attendance. We find that it does: the higher the fraction of peers that were ex post "justified" for failed attendance in the previous one or two months, the lower the child's own attendance in a given month. This suggests that program beneficiaries extract signals of lax enforcement from other people's experience and respond accordingly.

It should be noted that we use several different strategies to be able to interpret a negative correlation between peer warnings and own attendance decisions as an effect of learning about the quality of enforcement. In particular, we control for direct peer effects and correlated shocks by controlling for the attendance of an individual's peers and analyze if there is an additional effect of peers receiving warnings on the individual's attendance. Furthermore, we make use of the fact that we do not only observe the fraction of the individual's peers who receive warnings, but also the fraction of peers of each of the individual's siblings. Lastly, we make use of signals of implementation quality that are decided at different institutional levels: warnings (and loss of benefit) are implemented at the federal level, while ex-post justifications of a child's failure to comply is in the hand of the school the child attends. This allows us to disentangle a potential role of peer effects (at the classroom, grade or school level and between siblings) from learning effects.

This paper is related to a large literature on conditional cash transfer programs (on Progresa/Oportunidades see, e.g., Attanasio, Meghir, and Santiago (2011), Attanasio, Meghir, and Szekely (2003), DeBrauw and Hoddinott (2010), Janvry and Sadoulet (2006) and Todd and Wolpin (2006), on Bolsa Escola (the predecessor program of "Bolsa Familia") see, e.g. Bourguignon, Ferreira, and Leite (2003), Bursztyn and Coffman (2010), de Janvry, Finan, and Sadoulet (2009) and on Bolsa Familia see a recent evaluation study by IFPRI (Hoddinott and others, see press release).[1] In terms of literature on schooling in Latin America more generally, for example Harbison and Hanushek (1992) discuss programs to improve school quality of the poor in the Northeast of Brazil.

Our paper is also related to the literature on peer effects and social interactions, even though our focus is not to identify conventional or "direct" peer effects (i.e. the effect of non-attendance of peers on own attendance), but instead we are interested in identifying the effect of peers receiving warnings (as a signal about the quality of enforcement) on an individual's attendance decisions. One example of a paper that is directly interested in the role of information and social interactions, but in the context of retirement plan decisions, is Duflo and Saez (2003). Another related paper on learning from peers in the context of technical change in agriculture is Foster and Rosenzweig (1995). Some examples of the related literature on peer effects and social interactions more generally are Aizer and Currie (2004), Banerjee, Chandrasekha, Duflo, and Jackson (2010), Ding and Lehrer (2006), Duflo, Dupas, and Kremer (2010), Hanushek, Kain, Markman, and Rivkin (2003), Hoxby (2000) and Hoxby and Weingarth (2005). Two survey articles that provide an overview and synthesis of the literature on how social networks influence behaviors are Jackson and Yariv (2010) and Blume, Brock, Durlauf, and Ioannides (2010).

The following papers analyze peer effects in the context of conditional cash transfer programs in Colombia and Mexico: Barrera-Osorio, Bertrand, Linden, and Perez-Calle (2011) analyze a randomized CCT in Colombia with three different treatments based on school attendance and randomize on the student level, which allows them to generate intra-family and peer-network variation. They find evidence that siblings (particularly sisters) of treated students work more and attend school less than students in families that received no treatment. The authors also find that indirect peer influences are relatively strong in attendance decisions with the average magnitude similar to that of the direct effect. Other papers that analyze peer effects and social interactions in the context of Progresa/Oportunidades are Angelucci, Giorgi, Rangel, and Rasul (2010) on family networks, Angelucci and Giorgi (2009) on consumption spillovers and Bobonis and Finan (2009) on peer effects in schooling.

This paper is indirectly related to the tax enforcement and crime deterrence literature that have looked at the role of enforcement. Aside from the difference in the context (tax evasion and crime constitute breaches of the law, while failure to meet welfare program conditions does not), to the best of our knowledge the above literature has not provided sound empirical evidence of multiplier effects, that is peer learning about enforcement.

[1] Our paper is also related to a paper by Dee (2011) on conditional cash penalties, as Bolsa Familia is a transfer that families usually receive unless they do not comply to the conditionalities, in which case they get punished with increasing degree of severity. In the case of Progresa/Oportunidades on the other hand a family receives a transfer per child that attended in the previous month.

To mention one example in the tax enforcement literature, Pomeranz (2011) analyzes the role of information in tax enforcement. In particular she tests claims that the value-added tax facilitates tax enforcement by generating a paper trail on transactions between firms. Pomeranz (2011) finds that the paper trail leads to spillovers in that the impact of a random audit announcement is transmitted up the VAT chain, increasing compliance by firms' suppliers.

Two papers in the crime deterrence literature that investigate the effects of learning about arrest probabilities on criminal behavior are Sah (1991) and Lochner (2007) among others. Sah (1991) provides a theoretical analysis of crime based on a model in which individual beliefs about the probability of punishment are determined by the number of people they observe committing crime and their arrest rates. His theory suggests interesting dynamic responses to changes in criminal enforcement policy as well as levels of segregation. Lochner (2007) analyzes the effect of own and siblings' criminal history and arrests on perceptions about arrest probability (using data on people's subjective beliefs) and estimates the effect of perceptions about arrest probabilities on criminal behavior in a structural model.

2 Background Information on the Bolsa Familia Program

The Bolsa Familia Program reaches around 11 million Brazilian families, that is 46 million poor people (equivalent to 25% of the Brazilian population) with a budget of over 12 billion reais (USD 6 billion). Thus Bolsa Familia reaches nearly three times as many people and is about three times as large in terms of budget as the well-known conditional cash transfer program Progresa/Oportunidades.

The Bolsa Familia Program (BFP) results from consolidating four different programs (Federal Bolsa Escola Program, Bolsa Alimentacao, Auxilio Gas, Fome Zero) into one single program (see, for example, World Bank Discussion Paper (2007)). It was launched in 2003 by Lula. The election-free year 2005 was used as year of consolidation to strengthen the core architecture of the program (20 legal and operational instruments were issued, institutionalizing various aspects of the program and its decentralized implementation) and to strengthen the program's registry of families (Cadastro Unico) and the monitoring of conditionalities. In 2006, the Ministry of Social Development embarked on initiatives to promote further vertical integration with sub-national CCTs and integrated the conditional transfers paid under the Child Labor Eradication Program (PETI) into the BFP. The Ministry started providing incentives for quality management and rewarding innovations in the decentralized management of the BFP. The agenda for 2007 and beyond was to further strengthen the basic architecture of the program, particularly in the areas of monitoring and verification of conditionalities, to strengthen the oversight and control mechanisms and continued improvements on the program's targeting system.

The targeting of the program was conducted in two steps. The first step consisted of geographic targeting on federal and municipality level. The federal government allocated BFP quotas to municipalities according to estimates of poverty. [2] Within municipalities spatial maps of poverty were

[2] Original municipal level allocations were established by comparing eligibility criteria (per capita income thresholds of R$ 100) to the 2001 national HH survey (PNAD) combined with 2000 Census. In 2006 MDS adjusted overall targets as well as specific program quotas for municipalities using PNAD 2004. Revisions resulted in adjustments of specific municipal quotas with some municipalities facing a reduction in BFP allocation (only few municipalities had beneficiary totals that were above the newly established quotas, but nobody was cut out of the program).

used to identify and target geographic concentrations of the poor. The second step was to determine eligibility at the household level. Eligibility is determined centrally by the Ministry of Social Development (MDS) based on household registry data that is collected locally and transmitted into a central database known as the Cadastro Unico.

BFP provides two types of benefits: basic and variable, according to family composition and income. Families with a monthly per capita family income of up to R$60 (US$30) are classified as "extremely poor", families with between R$60 and R$120 are classified as "moderately poor". The base benefit is provided only to families in extreme poverty, regardless of their demographic composition. Both extremely poor and moderately poor families receive a variable benefit which depends on the number of children in the family (capped at three to avoid promoting fertility) and on whether the mother is pregnant or breast-feeding.

The benefit amounts are as follows in the period of analysis, 2008 and 2009:

- January 2008 to June 2008: base benefit R$ 58 (US$ 30 and equal to the per capita income threshold for the extremely poor), variable benefit R$ 18.
- June 2008 to July 2009: base benefit R$ 62, variable benefit R$ 20.
- August 2009 to March 2011: base benefit R$ 68, variable benefit R$ 22.

To illustrate the magnitude of the program transfers, imagine a family with three children that earns just as much to still be classified as extremely poor (i.e. they have a monthly per capita income of R$60 and a family income of R$ 300). This family will receive monthly transfers of around R$120, which amounts to 40% of their total family income.

The BFP cash transfers are conditional on all age-relevant family members complying with school attendance "conditionalities". Each school-aged child has to attend at least 85% of days each month. This element of "joint responsibility" for children in the same family is a quite unique feature of BFP (compared to other well-known CCTS such as Progresa/Oportunidades).

Bolsa Familia conditionalities have been widely publicized in Brazil, and are spelled out in a booklet issued to each beneficiary family (Agenda de Compromissos).

The consequences for non-compliance in the different stages of conditionaliy are as follows. In the first case of non-compliance the family receives a warning without financial repercussion. With the second warning, benefits are blocked for 30 days after which the family receives the accumulated benefit of the previous and the current month. The third and fourth warnings lead to a loss of benefits for 60 days each time. After the fifth warning, the benefit is canceled and the family loses eligibility (according to the general rules, the family can reapply to the program 18 months later, for exceptions see description below). The time at which families receive warnings is when they withdraw their transfer money at the bank.

The implementation of this conditionality scheme involves different actors. First of all, children's attendance is recorded by the school teachers. The school sends the attendance lists of students to the municipality: they report the exact fraction of days attended in case attendance was below 85%, otherwise they only report that the student complied. Each municipality collects the lists and sends them to the Ministry of Social Development (MDS), which determines whether a family complied or not in a given month, i.e. whether all the children between 6 and 15 have attended at least 85% of days. In case of noncompliance, MDS establishes which warning the family should receive and whether the family is entitled to the transfer for that month (based on the warning stage reached). The MDS sends this information to the Caixa, which transfers the benefit amount to the bank account of the family if the family is entitled to receiving the transfer for that month.[3]

The family can withdraw their transfer money starting at a pre-specified date each month with a Bolsa Familia "bank card". At the time of withdrawal, the family receives a message, which declares the warning stage, the month of failure to which the warning refers, the names of the child(ren) who failed and which type of warning they might receive in the next instance of non-compliance (for example "if you fail to comply again, your money might be suspended").

The following exemptions from the standard procedure can be observed. Firstly, the school can justify insufficient attendance in a given month ex-post (which we observe in the data). In that case the child will count as having complied with the rules in that month. Secondly, while according to the general rule a family who receives the fifth warning has to leave the program and can only be admitted back after 18 months, the municipality has the right to readmit the family more quickly in case they start complying after the fifth warning.

3 Data

We make use of administrative data on the program from the Brazilian Ministry of Social Development (MDS), which contains the following information: The household registry data (Cadastro Unico) contains extensive information on families' background characteristics and on each individual member of the household, such as age, gender, race, marital status, education, employment status and occupation of each adult household member, per capita expenditures, ownership of durable goods etc and schooling history of each child.

We have information on monthly school attendance of each child for 2008 and 2009 (we know the exact fraction of days failed below the threshold of 85% and otherwise we only know that the condition was met), monthly information on warnings (so we can create the complete warning history of each family) and monthly information on benefits that the family received (or should have received) and whether the benefit was blocked or suspended in a given month.

A third source of data is the School Census, which contains information on all children who are enrolled in a given school. We merge the administrative data described above with School Census data (merge based on either social security number of the child –if available– and otherwise based on area code, school code, grade, full name and birth date of child), so that we can identify the (BFP) peers who are in the same school, grade and class for each child in the family. Also this data provide information on whether a child passed or failed a given grade.

Table 1 presents summary statistics of the variables used in our analysis.

[3]The Caixa Econmica Federal has been contracted as BFPs operating agent. The Caixa is a federal savings/credit union organization, which –apart from banking services– has traditionally provided payments issuance services for federal assistance programs, because of its' broad network which guarantee its' presence in all Brazilian municipalities (it operates over 2,000 agencies nationwide, and is linked with close to 9,000 lottery points and over 2,000 banking correspondents). The Caixa consolidates and manages the national registry database for social programs, the Cadastro nico, assigns registered individuals the unique Social Identification Number (NIS), and makes payments directly, crediting beneficiaries electronic benefit cards (EBCs) on a monthly basis through its extensive banking network. The Caixa also designed and operates the software currently used by the Ministry of Education (MEC) for consolidating the information resulting from the monitoring of compliance with conditionalities.

4 Empirical strategy

4.1 Response to Private Signals of Enforcement

The first goal of this paper is to analyze how families respond to receiving warnings. The outcome variable we use in the following analysis is whether a child fails to comply in a given month (i.e. attends less than 85% of days) or not. For ease of exposition, we refer to this variable as "failure" and we study the probability to "fail", but it should be clear that this term is not used in the traditional meaning of grade repetition, but rather it refers to failure to comply with the conditionality embedded in BFP.

We analyze how families respond to receiving different levels of warnings (warning levels 1 to 5, as described in the previous section), how their response depends on how much money they lose and how responses differ depending on whether the family is likely to be liquidity constrained or not.

One challenge in the identification of the effects of warnings stems from the fact that some families have a much higher propensity to fail and they are the ones who reach high warning stages. This is likely to lead to a positive correlation between "failure" to comply and warnings received when using cross-sectional data.

We address this problem using family fixed effects in the rest of the analysis to control for time-invariant differences in the propensity to fail. This implies that we use variation in warnings (and benefit amounts etc) within family over time to identify their effect on attendance behavior each month.

Our baseline specification for estimating the effects of own warnings is:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^k + \gamma X_{iht} + D_t + D_h + \epsilon_{iht} \tag{1}$$

where i denotes the child, h the household, t the month, Y is a dummy equal to 1 when the child attends less than 85 percent of the school days in a given month ("failure"); WS^k denotes a dummy equal to one if the household in in warning stage k (with $k = 1, ..., 5$), X is a vector of child level controls including gender, age, number of brothers and sisters in different age brackets (0-5, 6-10, 11-15, 16-18); D_t denotes month and year fixed effects to control for example for seasonality effects; D_h denotes household fixed effects and ϵ is the error term. We estimate (1) using a linear probability model (robustness results obtained estimating a probit will be included in the Appendix of the next version of this paper).

We expect the coefficients α_k to be negative and increasing with k in absolute value, as later warning stages correspond to increasingly severe punishments (see background section). It is important to keep in mind that only subsets of families reach the different warning stages. Thus the coefficients α_k, $k = 1, ..., 5$ are identified through different subsets of families. For this reason we will also compare the coefficients of the different warning stages separately for the different subsets of families.

We also test whether the effect of receiving a given level of warning differs depending on whether the household is liquidity constrained. This is done by adding in (1) a set of interaction terms $WS_{ht}^k * LIQ_h$, where LIQ_h is our proxy for liquidity constrained households, i.e. a dummy equal to one if the household does not own the dwelling where they live. While admittedly imperfect, this proxy is the best we can use given the very limited information contained in the household registry of BFP beneficiaries.

We then analyze how the likelihood to fail depends on the potential loss that the family faces, i.e. the benefit amount for which the family is eligible. For this purpose, we make use of exogenous changes in benefit amounts originating from two sources. The first is a policy-induced change which occurred at two points in time during our sample period: in July 2008 (announced in June 2008, an increase of 8%) and in August 2009 (announced in July 2008, an increase of 10%). The second source of variation is mechanically induced by the demographic structure of the household when children transition across age ranges that entitle the household to different benefit amounts.

Finally, we test whether the quality of enforcement matters, for a given warning stage. As proxies for the quality of enforcement that a household experiences with respect to its own warnings, we employ two measures. The first is the delay with which a given warning is received, compared to the month in which the conditionality on attendance was not met. We estimate:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^k + \sum_{k=1}^{5} \beta_k WS_{ht}^k * DEL_h^k + \gamma X_{iht} + D_t + D_h + \epsilon_{iht} \tag{2}$$

where DEL_h^k is the number of months (in deviation from the median of 5 months) that passed between the month in which the attendance condition was not met and the month in which household h was notified the punishment in the form of transitioning to warning stage k. We expect $\beta_k > 0$, i.e. delays in enforcement should be associated with a relatively higher probability of failing to attend school, compared to timely enforcement. In other words, warnings received with longer delay are less effective.

We also consider an “extreme”form of lax enforcement, i.e. the ex post justification given by the school when a child failed to attend, so that the household does not advance in warning stage. We estimate:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^k + \sum_{n=1}^{3} \beta_n JUST_{h,t-n} + \gamma X_{iht} + D_t + D_h + \epsilon_{iht} \tag{3}$$

where $JUST_{h,t-n}$ is a dummy equal to 1 if household h has been justified in month $t-n$, with n going from 1 to 3, i.e. we consider three lags of this variable. Again, we $\beta_n > 0$, i.e. to the extent that justifications are interpreted as signals that failure to comply with conditionality is not punished, justifications should be associated with a relatively higher probability of failing to attend school.

4.2 Response to Public Signals of Enforcement

Families can not only learn from their own private signals of enforcement (e.g., if they receive warnings), but also from “public signals,” that is from the experience of other children who receive warnings, lose their transfer or even have to leave the program. In particular, children who are in the same class, grade or school as the family’s own children are likely to be key sources of information regarding the implementation of the program. We would expect the family to decide about attendance in each month based on this information. To discuss our empirical strategy it is useful to start from a simple benchmark specification which is not exactly the one we estimate (see below) but which helps us highlight identification challenges:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^k + \beta \ PEERWARN_{iht} + \gamma X_{iht} + D_t + D_h + \epsilon_{iht} \quad (4)$$

where $PEERWARN_{iht}$ denotes several possible indicators for the presence of peers who receive a warning in month t. As a preview, we find a significant negative correlation between own failure to attend in a given month and the fraction of peers who are warned ($\widehat{\beta} < 0$), that is, the likelihood to fail decreases when peers receive warnings. Also we find a significant positive correlation between own failure to attend and the fraction of peers whose failure is justified ex-post.

We need to address several important identification challenges before we can interpret this negative correlation as learning. In what follows we group these identification challenges under two headings: correlated shocks and "conventional" peer effects.

Correlated shocks

The first threat to identification are correlated shocks that may directly affect a student and her peers, thus inducing a correlation between $PEERWARN_{iht}$ and ϵ_{iht} in equation (4). Consider for example an economic shock leading to an increase in the opportunity cost of schooling (e.g., due to labor market conditions) in the area where individual i lives. In response to such a shock, both individual i and her peers would be more likely to fail and thus more likely to receive a warning. If the shock is persistent, the shock would still induce individual i and her peers to fail more several months later when the warning is received. However, this type of mechanism would generate a *positive*, not negative, correlation between one's failure and peers' warnings ($\widehat{\beta} > 0$), while what we find in the data is a negative correlation ($\widehat{\beta} < 0$). This negative correlation is consistent with a "scare"effect, i.e. student i decreases her failure to avoid the sanctions that her peers experienced, but not with positively correlated economic shocks. The shock could of course be mean reverting, which is a concern that we address below.

A different type of correlated shock may be linked to changes in school policies, e.g. the teacher or headmaster become stricter in registering students' non-attendance. This would lead to a higher number of peers receiving warnings. At the same time, everyone would fail less because it is known that the teacher or headmaster have become stricter, not because they extract signals on enforcement from peers' warnings. This type of correlated shock would thus generate an estimate with the same sign as we find, i.e. ($\widehat{\beta} < 0$).

To address this concern, as a robustness check we include the peer warnings as a lead variable, i.e. we include $PEERWARN_{ih,t+1}$ in equation (4). Our reasoning exploits the time lag from the moment in which a failure in attendance occurs and the moment in which the student gets warned, which has a median value of 5 months. If children decrease their failure because of greater strictness, they will start attending more soon after they experience this greater strictness, which would be before the time in which warnings are received (e.g., before 5 months). This means that changes in attendance would precede warnings, so we should find a negative and significant coefficient on $PEERWARN_{ih,t+1}$. If not, then it is unlikely that the greater strictness that started 5 months ago (median delay between failure and warning), has an effect exactly 5 months afterwards (when peers receive the warning) but not, say, 4 months afterwards. This strategy can also address the concern of mean reverting shocks: Again it is highly unlikely that a shock that increased failure 5 months ago, suddenly decreases failure with the exact same delay as the warning happens (but has no increasing effect on failure 4 months after the initial shock).

To summarize, the augmented version of equation (4) that we estimate is:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \beta \sum_{n=-1}^{+1} PEERWARN_{ih,t+n} + \zeta \sum_{n=-1}^{+1} PEERFAIL_{ih,t+n} + \gamma X_{iht} + D_t + D_h + \epsilon_{iht} \tag{5}$$

Conventional peer effects

The second concern is related to "conventional" peer effects. The mechanism would be as follows: A shock of some sort (economic shock or school level shock) increases children's likelihood to fail. Individual i's peers receive warnings and in response to their own warnings they fail less. As a consequence also individual i fails less because she observes her peers failing less and not because she extracts signals on enforcement from peers' warnings. This is what we call a "direct" or "conventional" peer effect as opposed to learning about the strictness enforcement from peers' warnings.[4] To address this concern, we proceed with four strategies.

First, we can control directly for the fraction of peers (in i's class) who fail, to analyze if the warning of i's peers has an independent effect on individual i's likelihood to fail. This makes it difficult for us to find any effect of peer warnings for the following reason. This strategy implies that we compare two classes with the same fraction of students failing, while in one class the fraction of children warned is larger. If people respond to own warnings, then children in the class with more warnings need to have a general tendency to fail *more* than the children in the class with fewer warnings. To see a learning effect, the individual should fail *less* in a class with more warnings, while –according to this argument– she is in a class where people are prone to fail *more*.

Second, we exploit warnings received not by individual i's own peers, but by the peers of i's siblings. In addition to the fraction of individual i's own peers who are warned, we also include the maximum of the fractions of i's siblings' peers who got warned, as in the following specification

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \beta_1 PEERWARN_{iht} + \beta_2 MaxPEERWARN_{-i,ht} + \gamma X_{iht} + D_t + D_h + \epsilon_{iht} \tag{6}$$

where the subscript $-i$ indicates that the Max is taken at the household level excluding individual i's own peers. Then we compare β_1 and β_2 for two sub-samples: in the first, $PEERWARN_{iht} \geq MaxPEERWARN_{-i,ht}$, i.e. the maximum fraction of peers warned among all children is that if i's own peers. In the second sub-sample $PEERWARN_{iht} \leq MaxPEERWARN_{-i,ht}$, i.e. the maximum is obtained by one of i's siblings. Our learning story is consistent with a case where $|\beta_1| > |\beta_2|$ in the first sub-sample, while $|\beta_1| < |\beta_2|$ in the second. In other words if warnings matter because they signal enforcement, then the "toughest" signal is the one that should affect attendance the most. On the other hand, this pattern could not be generated by a pure "conventional peer effect" (i.e. lower failure of children in i's own class) because such mechanism would imply $|\beta_1| > |\beta_2|$ in both sub-samples. In particular, direct classroom peer effects could not lead to a significant effect of the warnings received by individual i's siblings' peers, who are typically in different grades or schools.

[4]The direct peer effect could be due to preferences for school attendance that depend on the presence of peers, but also be due to interpreting peers' attendance as a signal that attendance is in some way beneficial (for example due to high returns to schooling or due to strict enforcement and possible loss of benefits in case of non-attendance. To identify learning effects in a clean way, we control for direct peer effects in the first set of results, which of course might lead to an underestimate of the "true" learning effects, if part of the learning about enforcement happens through observing ones' peers attend more.

Third, we might still be concerned about peer effects between siblings that could lead to the pattern discussed above as an alternative explanation to learning effects. In particular, if a large fraction of the peers of i's siblings receive warnings, then i's siblings might attend more (either because of the warning or due to peer effects) and then i might decide to attend more because her siblings attend more. The mechanism underlying this direct peer influence among siblings may be one of role models, or it may be related to time use (for example, when two brothers usually play soccer together and if one goes to school the other one loses his soccer partner). It seems plausible that this type of peer effect between siblings should be stronger, the closer the siblings are in age and more likely if the two siblings are of the same gender.

To analyze the importance of this alternative explanation, we repeat the last test comparing the effect of fraction of own peers warned and of the maximum fraction of siblings' peers warned who are in schools *different* from i's school. In other words, we estimate a modified version of (6) where the variable $MaxPEERWARN_{-i,ht}$ is constructed from siblings who attend a different school than i. Siblings usually go to different schools only if they differ enough in age and thus have to go to different types of schools (such as elementary, junior high school etc). If the learning story is true, then the "maximum fraction of siblings' peers warned who are in schools different to i's school" should have an effect that is as strong as the "maximum fraction of all siblings". If on the other hand, peer effects between siblings are driving the results, then we would expect the former to have a smaller effect, because in that case the siblings are further apart in age and thus less likely to be playmates.

This leads us to a closely related test, based on the idea that children of the same gender are also more likely to be playmates. Thus, we test if there is a differential effect of the "maximum fraction among the siblings" depending on whether the maximum is taken among siblings of the same gender or the opposite gender. If the learning story is true, the maximum should have the same effect independent on whether it is based on siblings of same or opposite sex. If the sibling peer effect story is true on the other hand, we would expect the maximum among opposite-sex siblings to have a smaller effect (as these are less likely to be playmates).

Fourth, to lend further credibility to the hypothesis that people learn from their peers and use an additional strategy to address the concern of peer effects between siblings, we make use of the following feature of the implementation of BFP: Warnings and ex-post justifications of non-compliance are implemented at different institutional levels. The implementation of warnings (and loss of benefits) is in the hand of the Ministry of Social Development, i.e. done at the federal level. The ex-post justification of a child that failed to comply in a given month is in the hand of the school (teacher or headmaster). If the effects that we find are due to learning effects, then the attendance behavior of an individual should be affected by warnings of her siblings' peers who are in different schools (because it is a signal of enforcement quality at the federal level), but should not be affected by justifications of her siblings' peers who are in different schools (or less affected, if justifications are correlated across schools). Instead both warnings and justifications should affect an individual's behavior, if they happen to peers who are in the same school (class) as the individual. If on the other hand, the effects we find are peer effects among siblings, then in both cases (with warnings and justifications) the peers of the siblings would have an effect on the individual via "sibling" peer effects, i.e. effects of warnings and justifications should be symmetric.

5 Analysis

5.1 Response to Private Signals of Enforcement

In this section we analyze how families respond to receiving warnings. Our outcome variable, “failure”, is a dummy equal to one if a child fails to comply in a given month, i.e. attends less than 85 percent of the days, and 0 otherwise.

We analyze how families respond to receiving different levels of warnings (warning levels 1 to 5, as described in the previous section), how their response depends on how much money they lose and how responses differ depending on whether the family is likely to be liquidity constrained or not (proxied by whether the family owns the dwelling they live in or not).

[Insert Table 2]

As discussed in the previous section, estimating the effects of warnings on attendance is complicated by the fact that some families have a much higher propensity to fail and they are the ones who reach high warning stages, which leads to a positive correlation between ”failure” to comply and warnings received when using cross-sectional data (see specification (1) in Table 2).

We address this problem using family fixed effects in the rest of the analysis to control for differences in the propensity to fail and thus we make use of variation in warnings (and benefit amounts etc) within family over time to identify their effect on attendance behavior each month. Table 2 shows that using family fixed effects, the likelihood of non-attendance (that is failure to comply in a given month) decreases when the family receives a warning. This negative effect is of increasing magnitude when the warning stage increases, as would be expected given the gradual increase in the severity of the punishment (see specifications (2)-(5) in Table 2):

A preliminary consideration can be useful to interpret the effect size. When a family reaches a given warning stage, say WS^k their reaction should depend on what happens if they do not comply in the future months, i.e. on what would be the penalty associated with WS^{k+1}. In column 2 of Table the coefficient $-.017$ on Warning Stage 1 implies that when a family moves to the first warning stage (which does not lead to any loss of benefit) they are 1.7 percentage points less likely to fail to comply with attendance. This negative effect is consistent with the fact that in the next case of non-compliance the benefit will be blocked for 30 days (but received later), which is costly for poor families. Moving to the second warning stage has a significantly larger effect than the first warning (coefficient $-.037$). This can be explained by the fact that in an additional instance of non-compliance, the benefit will be suspended for 60 days and the money lost. The coefficient on Warning Stage 3 is slightly (but not significantly) larger at $-.048$, as the penalty for Warning Stage 4 is the same as for Warning Stage 3. The coefficient on the fourth warning stage, on the other hand, is $-.083$ and is significantly larger than those of all previous warning stages: the family cannot fail anymore without the threat of losing eligibility. In the light of the general rules, it is surprising that even the fifth warning has a large and negative effect on failure, though this coefficient is not significant in other specifications (see next tables). One possible conjecture to explain this result is that families react because they know that municipalities have the right (and make use of the right) to readmit beneficiaries in case they start complying after the fifth warning.

In columns (3) and (4) of Table 2 we analyze if the likelihood to fail is smaller when the “potential loss” (benefit amount for which the family is eligible) is larger. The variable “potential loss” has a mean of R$ 107 and a standard deviation of R$ 27. To identify the effect of a change in potential loss while controlling for family fixed effects, we make use of two sources of exogenous changes: one comes from changes in the ages of children that become or cease to be eligible for the program; the other comes from government mandated changes in benefit amounts in July 2008 (announced in June 2008) by 8 percent and in August 2009 (announced in July 2008) by 10 percent.

When we include variable potential loss on its own, its coefficient is negative but not significant at conventional levels. We interact the potential loss with the new warning stage reached to test if a given monetary amount has a larger effect in warning stages that correspond to more severe punishments (e.g., actual loss of money versus blockage with deferred payment). To be precise, we interact the warning level with a rescaled variable of potential loss (rescaled by the median potential loss) to be able to interpret the effect of the warnings stages at the median level of families' potential loss (because no family has a potential loss of zero). We find that the effect of the first warning does not depend on the potential loss (the sum of the two coefficients on the main effect and on the interaction of warning stage 1 with potential loss is not significantly different from zero). In the case of the third warning stage on the other hand, the effect is significantly larger, the larger the potential loss. It is difficult to interpret a significant positive coefficient on the interaction of fifth warning stage with potential loss, but it is important to stress that very few families actually reach this warning level.

As mentioned above, in the second instance of non-compliance money will be blocked, which is costly for poor families in general, but even more so for families that are particularly cash constrained. This implies that the first warning should have a particularly strong effect on families who are liquidity constrained, because they would suffer most from a blockage of their transfer in case of an additional failure. To proxy for being liquidity constrained we make use of whether the family owns the dwelling they live in. Specification (5) of Table 2 shows that families who are most likely to be liquidity constrained, respond particularly strongly to the first warning (the coefficient is negative also on higher warning levels but not significant).

[Insert Table 3]

We now move to an analysis of whether the "quality of enforcement" matters for the response. First, we analyze if the effect of warnings depends on the delay with which the warning was received (i.e. how many months after the month of non-compliance). The median delay time is 5 months, but can vary between 2 and 10 months. Table 3 shows that the effect of a warning is significantly smaller, if the warning was received with a longer delay. This is true when we interact warnings with delay in months rescaled by the median (column 1) and also when we interact warnings with a dummy for whether the length of delay was above the 80th percentile (column 2).

The second measure of "quality of enforcement" that we use is whether a child's non-attendance in a given month gets justified ex-post by the school (see background section for details), in which case the non-attendance is not counted as non-compliance. Specification (3) in Table 3 shows that the likelihood of failure is higher, if a child's non-attendance was justified ex-post in the previous month. Of course, this could be due to serially correlated shocks that affect the attendance behavior in two consecutive months. For this reason, we investigate further whether there is a signal of quality of enforcement in these "justifications" in the next section, when analyzing if families learn from justifications of the children's peer and respond in their attendance behavior.

5.2 Response to Public Signals of Enforcement

In this section we examine how families react to the signals of enforcement received not directly by them but by the "peers"of their children. We use the following peer group definitions: school, grade and class of each child and we analyze two types of signals: Peers receiving any warnings and peers getting justified ex post for their failure. The former is a signal of enforcement, while the latter is a signal of non-enforcement.

We construct two types of peer group variables: one is the fraction of peers in a class who were warned; the other is a dummy equal to one if at least one peer was warned. We also compare the effect of own peers being warned versus siblings' peers being warned.

First we discuss the results related to peers receiving warnings (Tables 4 and 5). Second, we discuss the results related to peers who receive justifications (Tables 6 and 7).

5.2.1 Peers Receiving Warnings

In column 1 of Table 4 we include the "fraction of own peers who receive any warning" as a main effect and a lagged effect in addition to own warnings. We include a lagged variable of peer warnings, because warnings happen in the last two weeks of a month and thus it is ex ante difficult to establish whether children learn about their peers' warnings in time to adjust their attendance in the current month or in the following month. The coefficient on the fraction of peers warned in column (1) is $-.014$, significant at the 5 percent level, indicating that there is a significant negative correlation between own peers receiving a warning in a month and own attendance in the same month. The coefficient on the lag is instead zero.

To address the concern that this result might be due to direct peer effects, in column (2) we control for the fraction of peers who fail to attend in that month. We find that the fraction of peers who fail is significantly positively correlated with own failure (which could be due to peer effects or common shocks). At the same time we still find a highly significant and even larger negative effect of peers' warnings on own failure after controlling for peers' failure (coefficient $-.026$), which is consistent with the interpretation that people learn about stricter enforcement (and reduce their failure in response) when observing peers being warned.

In specification (3) of Table 4, we include the fraction of peers warned as lead variable to rule out the possibility that our result is driven by changes in school policies, e.g. the teacher or headmaster becoming stricter in registering students' non-attendance (and also to address the concern of mean reverting economic shocks). As explained in section 4.2, if children decrease their failure because of greater strictness, they should start attending more soon after they experience this greater strictness, which would be before the time in which warnings are received (the median lag being 5 months). This means that we should find a negative and significant coefficient on the lead variable. In column 3, we find that the coefficient on the lead variable is negative but smaller than the main effect and not significant, which is consistent with our learning interpretation.

In the last specification of Table 4, we use an alternative variable to capture the effect of peer warnings. In principle it is not clear that there should be a linear relationship between the fraction of peers warned and own attendance. It may be the case that what matters is whether at least one peer receives a warning, which is what we test in specification (4). We again find a negative coefficient on this variable, but the significance level is reduced to 10 percent. So the number or fraction of peers who are warned affects the quality of the signal, which is to be expected.

[Insert Table 5]

In Table 5 we provide further evidence for our learning hypothesis and analyze additional specifications to rule out further alternative stories (in particular related to classroom and sibling peer effects). If results were driven by classroom conventional peer effects, that is children warned in individual i's class decrease their own failure and therefore i decreases her failure, then the warnings received by the peers of i's siblings (who are in a different class) should not have an effect on i's behavior. For this reason we include two "peer warning" variables: one is the fraction of own peers warned and the other is the "maximum of the fractions of peers warned among the siblings of individual i". We then divide the sample into two groups: in the first group, individual i's "own fraction" is the maximum fraction warned among all children in a family and in the second group the maximum is achieved by the siblings of individual i.

As expected, we find that the "larger" signal has a more important effect. In particular, in the second group where the siblings have the larger fraction of peers with warnings, the variable "maximum of the fractions of peers warned among the siblings of individual i" is negative and significant (column 2 in Table 5). It is also significantly different from the coefficient on "own peers' warnings" (which is close to zero and insignificant in this subsample). On the other hand, in the subsample where the own peer variable is the maximum, only the the coefficient on "own peers' warnings" is significant (column 1).

While this helps to address the concern that the results are driven by classroom peer effects, we might still be concerned that our result can be driven by peer effects between siblings, as discussed in the list of identification challenges. In particular, if a large fraction of the peers of i's siblings receive warnings, then i's siblings might attend more and then i might decide to attend more because her sibling attends more. We hypothesize that this type of peer effect between siblings should be stronger, the closer the siblings are in age and for siblings of the same gender.

To analyze the importance of this alternative explanation, we repeat the last test using the variable "maximum fraction of siblings' peers warned" but firstly focusing solely on siblings at different schools and secondly on siblings of the opposite sex. Siblings usually go to different schools only if they are different in age and thus have to go to different types of schools (such as elementary, junior high school etc). If the learning story is true, then the "maximum fraction of siblings' peers warned who are in schools different to i's school" (or using only siblings of opposite gender) should have an effect that is as strong as the maximum fraction of all siblings. If on the other hand, peer effects between siblings are driving the results, then we would expect the former to have a smaller effect, because in that case the siblings are further apart in age (or of opposite sex) and thus less likely to be playmates.

Specification (4) in Table 5 shows that the coefficient on "maximum fraction of siblings' peers warned in different schools" is significantly negative and of similar magnitude as the coefficient on the "maximum fraction among all siblings". The same is true for specification (6), which shows that focusing on the peers of siblings of opposite gender has basically the same effect as regarding the maximum of peers warned among all siblings.

This provides evidence against a story of peer effects between siblings (or peer effects between children in different classes and grades) and in favor of a learning story where families learn about the quality of enforcement from their children's peers. We provide further supportive evidence against a story of peer effects between siblings in the next section.

5.2.2 Peers Receiving Justification of Absence

In this section, we analyze the effect of having peers who receive an ex-post justification for their failure. It is interesting to see if people learn not only from positive signals of the quality of enforcement (i.e. seeing people receiving warnings), but also from negative signals, or signals of lax enforcement. Furthermore, we use this section to lend further credibility to the hypothesis that people learn from their peers.

In particular, we perform the following test. Warnings and ex-post justifications of non-compliance depend on different institutions: The implementation of warnings is in the hands of the Ministry of Social Development, i.e. done at the federal level. The ex-post justification of a child that failed to comply in a given month is in the hands of the school. If the effects that we find are due to learning effects, then the attendance behavior of an individual should be affected by warnings of her siblings' peers who are in different schools (signal of enforcement quality at the federal level), but should not be affected by justifications of her siblings' peers who are in different schools (or at least they should be affected to a lesser extent, if justifications are correlated across schools). Instead both warnings and justifications should affect an individual's behavior, if they happen to peers who are in the same school (class) as the individual. If on the other hand, the effects we find are peer effects among siblings, then in both cases (with warnings and justifications) the peers of the siblings would have an effect on the individual via her sibling's attendance.

[Insert Table 6]

Table 6 shows that in the case of justifications, only justifications of the individual's own peers significantly increase own failure (independently of whether the maximum of peers justified happened in the "own" class or in the siblings' class). If instead a sibling -who attends a different school- has peers who get justified, this even decreases the individual's failure (coefficient on second lag is significant), while according to a sibling peer effect there should be positive effect (individual A's sibling has peers who get justified, so the sibling fails more and -if both siblings play with each other- also individual i should fail more).

While the negative coefficient on the maximum fraction of siblings' peers warned is prima facie difficult to rationalize, one may think that there could be a strategic game between siblings. Imagine that the parents punish those children who fail and if a single child is responsible for the family getting warned (and possibly losing money) the punishment is particularly harsh. If the individual learns that her siblings' peers' failure gets justified, this is an indication that also her sibling might get justified ex-post in case of non-attendance. In that case, her own failure would be noticed more (as discussed in the background section, parents receive a warning message when withdrawing transfer money and the message lists which child failed to attend and when), so the individual decides to attend. For siblings who attend the same school as the individual, there are two competing effects: on the one hand, the effect discussed above of not wanting to be the one responsible for the family getting warned. On the other hand, having peers' who get justified in the same school (but different class) might contain some information about the likelihood to get justified oneself. This should lead to a coefficient that is less negative for the siblings' peers' justification, than for the siblings' peers' justifications for siblings at different schools. Comparing specification (2) and (4) shows that indeed the coefficient on "maximum fraction of siblings' peers in other school warned" is larger than the coefficient on "maximum fraction of all siblings".

[Insert Table 7]

In Table 7 we analyze the effect of having peers who receive an ex-post justification for their failure. If this serves as a signal of how strict the school is in terms of justifying non-attendance, we would expect an individual to fail more in response to observing peers in their own class (or school) to be justified. In contrast to the case of peer warnings (which signal the quality of enforcement of the federal government agency), justifications are done at the school level. Thus in this case we would not expect peers of siblings in other schools to have an effect, if the effects we identify are learning effects.

Table 7 shows an alternative way to provide evidence of the existence of peer learning by controlling for the fraction of peers who fail (to control for direct peer effects or shocks). While without controlling for the fraction of peers who fail, the fraction of peers who are warned is significantly positively correlated with own failure, the coefficients lose significance when adding this control variable. If we use the alternative measure "any of own peers justified" (instead of the exact fraction), then the coefficient remains positive and there is a significant positive effect of having any peer who was justified on own attendance, which is consistent with the results in Table 6.

6 Conclusion

In this paper we study the implementation of a large-scale conditional cash transfer program "Bolsa Familia" in Brazil, which conditions transfers to poor families on children's school attendance. We analyze how people learn about the quality of enforcement and how this affects their behavior and –as a consequence– the effectiveness of the program.

We find that individuals respond to incentives and finetune their behavior in response to private as well as public signals about the quality of enforcement of the program. (Non-)Enforcement does not only have a direct effect on the family affected, but also an important multiplier effect on other families, who learn from the experiences of their children's peers.

Our finding that people adjust their behavior to the strictness of enforcement implies that not only formal rules but actual enforcement of the program conditions can be crucial for the program effectiveness. This aspect might be particularly important for developing countries, as they might lack administrative capacity (or political will) to strictly enforce the rules. Thus the design of conditional welfare program should take into account this important dimension and be such that proper enforcement can be guaranteed or at least facilitated.

References

Aizer, A., and J. Currie (2004): "Networks of neighborhoods? Correlations in the use of publicly-funded maternity care in California," *Journal of Public Economics*.

Angelucci, M., and G. D. Giorgi (2009): "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 99(1), 486–508.

Angelucci, M., G. D. Giorgi, M. Rangel, and I. Rasul (2010): "Family Networks and School Enrollment: Evidence from a Randomized Social Experiment," *Journal of Public Economics*, 94(3-4), 197–221.

Attanasio, O. P., C. Meghir, and A. Santiago (2011): "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa," *Review of Economic Studies*.

Attanasio, O. P., C. Meghir, and M. Szekely (2003): "Using Randomized Experiments and Structural Models for Scaling-Up: Evidence from the Progresa Evaluation," *IFS Working Paper, EWP03/05*.

Banerjee, A., A. Chandrasekha, E. Duflo, and M. O. Jackson (2010): "The Diffusion of Microfinance in Rural India," *Working Paper*.

Barrera-Osorio, F., M. Bertrand, L. L. Linden, and F. Perez-Calle (2011): "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia," *American Economic Journal: Applied Economics*, 3(2), 167–195.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2010): *Identification of Social Interactions*. Handbook of Social Economics.

Bobonis, G. J., and F. Finan (2009): "Neighborhood Peer Effects in Secondary School Enrollment Decisions," *The Review of Economics and Statistics*, 91(4), 695716.

Bourguignon, F., F. Ferreira, and P. G. Leite (2003): "Conditional Cash Transfers, Schooling and Child Labor: Micro-Simulating Brazil's Bolsa Escola Program," *The World Bank Economic Review*, 17(2).

Bursztyn, L., and L. C. Coffman (2010): "The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian Favelas," *Working Paper*.

de Janvry, A., F. Finan, and E. Sadoulet (2009): "Local Electoral Incentives and Decentralized Program Performance," .

DeBrauw, A., and J. Hoddinott (2010): "Must conditional cash transfer programs be conditioned to be effective? The impact of conditioning transfers on school enrollment in Mexico," *Journal of Development Economics*.

Dee, T. (2011): "Conditional cash penalties in education: Evidence from the Learnfare experiment," *Economics of Education Review*, 30(5), 924–937.

Ding, W., and S. Lehrer (2006): "Do peers affect student achievement in China's secondary schools?," *Working Paper*.

Duflo, E., P. Dupas, and M. Kremer (2010): "Peer Effects and the Impacts of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review*.

Duflo, E., and E. Saez (2003): "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*.

Foster, A. D., and M. R. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103(6), 1176–1209.

Hanushek, E. A., J. F. Kain, J. M. Markman, and S. G. Rivkin (2003): "Does Peer Ability Affect Student Achievement," *Journal of Applied Econonics*, 18, 527–544.

Harbison, R. W., and E. A. Hanushek (1992): *Educational Performance of the Poor - Lessons from Rural Northeast Brazil*. Oxford University Press.

Hoxby, C. M. (2000): "Peer Effects in the Classroom: Learning from Gender and Race Variation," *NBER Working Paper 7867*.

Hoxby, C. M., and G. Weingarth (2005): "Taking Race out of the equation: School Reassignment and the Structure of Peer Effects," *NBER Working Paper*.

Jackson, M. O., and L. Yariv (2010): *Handbook of Social Economics*.chap. Diffusion, Strategic Interaction, and Social Structure.

Janvry, A. D., and E. Sadoulet (2006): "Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality," *The World Bank Economic Review*, 20(1).

Lochner, L. (2007): "Individual Perceptions of the Criminal Justice System," *American Economic Review*, 97(1), 444–460.

Pomeranz, D. D. (2011): "No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax," *Working Paper*.

Sah, R. (1991): "Social Osmosis and Patterns of Crime," *Journal of Political Economy*, 99(6), 12711295.

Todd, P. E., and K. I. Wolpin (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96(5).

Appendix

Table 1: Summary Statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
Male	1245129	0.5320	0.4990	0	1
Age	1245129	11.5700	2.7700	6	16
Nu of Sisters 6 to 10	1245129	0.3575	0.5809	0	4
Nu of Brothers 6 to 10	1245129	0.3974	0.6042	0	4
Nu of Sisters 11 to 14	1245129	0.3965	0.6073	0	4
Nu of Brothers 11 to 14	1245129	0.4456	0.6310	0	4
Nu of Sisters 15 to 17	1245129	0.2211	0.4583	0	3
Nu of Brothers 15 to 17	1245129	0.2525	0.4870	0	3
Rural	1245129	0.4115	0.4921	0	1
Failure to Attend	1245129	0.0363	0.1871	0	1
Warning Received	1245129	0.0426	0.2019	0	1
Benefit Amount	1245129	106.6800	27.0200	18	596
Warning Stage 1	1245129	0.3495	0.4768	0	1
Warning Stage 2	1245129	0.0821	0.2745	0	1
Warning Stage 3	1245129	0.0296	0.1694	0	1
Warning Stage 4	1245129	0.0080	0.0892	0	1
Warning Stage 5	1245129	0.0010	0.0320	0	1
Delay in Warning (in Months)	53042	4.0200	1.7500	2	10
Absence Justified	1245129	0.0080	0.0892	0	1
Nu of BFP Kids in Class	1245129	13.3300	5.2800	1	34
Nu of BFP Kids in Grade	1245129	47.5000	39.8200	1	284
Nu of BFP Kids in School	1245129	181.3800	139.0000	1	707
Nu of Kids in Class	1245129	28.3200	8.2600	1	78
Nu of Kids in Grade	1245129	110.8900	113.1600	1	981
Nu of Kids in School	1245129	530.8300	425.9700	3	2352
Frac of Peers Failed to Attend	1245129	0.0174	0.0593	0	0.8571
Frac of Peers Warned	1245129	0.0204	0.0718	0	1
Max of Sibs Peers Warned	1245129	0.0185	0.0696	0	1
Max of Sibs Peers Warned (Other School)	1245129	0.0080	0.0447	0	1
Max of Sibs Peers Warned (Sib Opp Sex)	1245129	0.0087	0.0488	0	1
Frac of Peers Justified	1245129	0.0047	0.0309	0	1
Max of Sibs Peers Justified	1245129	0.0037	0.0276	0	1
Max of Sibs Peers Justif (Other School)	1245129	0.0026	0.0232	0	1
Max of Sibs Peers Justif (Sib Opp Sex)	1245129	0.0022	0.0216	0	1
Frac of Sibs Justified	1245129	0.0099	0.0683	0	1
Frac of Sibs Justified (Other School)	1245129	0.0080	0.0617	0	1

Table 2: Effect of Own Warnings on Children's Attendance Behavior (Failure to Attend a Sufficient Number of Days in a Month).

Dependent Variable	**Decision of Noncompliance in Given Month**				
	(1)	(2)	(3)	(4)	(5)
Warning Stage 1	0.0089***	-0.0168***	-0.0151***	-0.0149***	-0.0160***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Warning Stage 2	0.0188***	-0.0368***	-0.0361***	-0.0362***	-0.0360***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Warning Stage 3	0.0271***	-0.0485***	-0.0468***	-0.0493***	-0.0488***
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Warning Stage 4	0.0447***	-0.0826***	-0.0828***	-0.0820***	-0.0813***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Warning Stage 5	0.0900**	-0.0914***	-0.0969***	-0.1182***	-0.0917***
	(0.041)	(0.026)	(0.024)	(0.026)	(0.03)
Potential Benefit			-0.0184	-0.0209*	
			(0.012)	(0.011)	
Warning St 1 * Pot Benefit				0.0670**	
				(0.032)	
Warning St 2 * Pot Benefit				0.0182	
				(0.09)	
Warning St 3 * Pot Benefit				-0.4546***	
				(0.16)	
Warning St 4 * Pot Benefit				-0.2505	
				(0.242)	
Warning St 5 * Pot Benefit				1.5902**	
				(0.737)	
Liq constrained					0.0027
					(0.002)
Warning St 1 * Liq constrained					-0.0052**
					(0.002)
Warning St 2 * Liq constrained					-0.0072
					(0.008)
Warning St 3 * Liq constrained					0.0064
					(0.014)
Warning St 4 * Liq constrained					-0.0272
					(0.031)
Warning St 5 * Liq constrained					-0.0009
					(0.04)
Controls	Yes	Yes	Yes	Yes	Yes
HH-FE	No	Yes	Yes	Yes	Yes
Observations	732298	1245129	1245129	1245129	1245129
R-squared	0.01	0.11	0.11	0.12	0.11

Notes: Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01. Included controls are age dummies, birth order dummies, month and year dummies and number of brothers and sisters in the age categories 0 to 5, 6 to 10, 11 to 15, and 16 to 18. All specifications (with exception of the first one) include family fixed effects.

Table 3: Effect of the Quality of Enforcement on Children's Attendance Behavior (Failure to Attend a Sufficient Number of Days in a Month).

Dependent Variable	**Decision of Noncompliance**		
	(1)	(2)	(3)
Warning Stage 1	-0.0195***	-0.0225***	-0.0720***
	(0.001)	(0.001)	(0.004)
Warning Stage 2	-0.0378***	-0.0434***	-0.1288***
	(0.003)	(0.003)	(0.008)
Warning Stage 3	-0.0572***	-0.0584***	-0.0642***
	(0.005)	(0.006)	(0.023)
Warning Stage 4	-0.0809***	-0.0935***	-0.0621
	(0.009)	(0.01)	(0.039)
Warning Stage 5	-0.1077***	-0.1009***	-0.1168
	(0.024)	(0.026)	(0.146)
Warning St 1 * Delay (in months)	0.0058***		
	(0.000)		
Warning St 2 * Delay (in months)	0.0063***		
	(0.001)		
Warning St 3 * Delay (in months)	0.0147***		
	(0.002)		
Warning St 4 * Delay (in months)	0.0205***		
	(0.006)		
Warning St 5 * Delay (in months)	-0.0097		
	(0.015)		
Warning St 1 * Long Delay (top 20%)		0.0251***	
		(0.001)	
Warning St 2 * Long Delay (top 20%)		0.0356***	
		(0.005)	
Warning St 3 * Long Delay (top 20%)		0.0420***	
		(0.009)	
Warning St 4 * Long Delay (top 20%)		0.0988**	
		(0.04)	
Warning St 5 * Long Delay (top 20%)		-0.0308	
		(0.024)	
Justification (lag 1)			0.1748***
			(0.009)
Justification (lag 2)			-0.0155*
			(0.009)
Justification (lag 3)			0.0046
			(0.008)
Controls	Yes	Yes	Yes
HH-FE	Yes	Yes	Yes
Observations	1245129	1245129	1245129
R-squared	0.12	0.11	0.25

Notes: Robust standard errors in parentheses (clustering at the household level). * p<0.1 ** p<0.05 *** p<0.01. Included controls are age dummies, birth order dummies, month and year dummies and number of brothers and sisters in the age categories 0 to 5, 6 to 10, 11 to 15, and 16 to 18. All specifications include family fixed effects.

Table 4: Effect of Own Peers Getting Warned on a Child's Attendance Behavior (Failure to Attend a Sufficient Number of Days in a Month).

Dependent Variable	**Decision of Noncompliance in Given Month**			
	(1)	(2)	(3)	(4)
Frac of Own Peers Warned	-0.0142**	-0.0260***	-0.0247**	
	(0.007)	(0.01)	(0.011)	
Frac of Own Peers Warned (lag)	-0.0016	-0.0066	-0.0021	
	(0.008)	(0.007)	(0.007)	
Frac of Own Peers Warned (lead)			-0.0103	
			(0.011)	
Any of Own Peers Warned				-0.0033*
				(0.002)
Any of Own Peers Warned (lag)				-0.0014
				(0.002)
Frac of Own Peers Fail		1.1380***	1.1366***	1.1387***
		(0.021)	(0.021)	(0.021)
Frac of Own Peers Fail (lag)		0.1175***	0.1019***	0.1158***
		(0.014)	(0.016)	(0.014)
Frac of Own Peers Fail (lead)			0.1038***	
			(0.013)	
Warning Stage 1	-0.0201***	-0.0202***	-0.0180***	-0.0205***
	(0.001)	(0.001)	(0.001)	(0.001)
Warning Stage 2	-0.0477***	-0.0475***	-0.0438***	-0.0479***
	(0.004)	(0.004)	(0.004)	(0.004)
Warning Stage 3	-0.0542***	-0.0587***	-0.0540***	-0.0588***
	(0.008)	(0.008)	(0.007)	(0.008)
Warning Stage 4	-0.1193***	-0.1022***	-0.1158***	-0.1022***
	(0.016)	(0.014)	(0.014)	(0.014)
Warning Stage 5	-0.0627	-0.0976	-0.1102	-0.0985
	(0.063)	(0.07)	(0.068)	(0.07)
Controls	Yes	Yes	Yes	Yes
HH-FE	Yes	Yes	Yes	Yes
Observations	1245129	1245129	1215057	1245129
R-Squared	0.12	0.23	0.24	0.24

Notes: Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01. Included controls are age dummies, birth order dummies, month and year dummies and number of brothers and sisters in the age categories 0 to 5, 6 to 10, 11 to 15, and 16 to 18. All specifications include family fixed effects.

Table 5: Effect of Siblings' Peers Getting Warned: All Siblings, Siblings at Other Schools and Siblings of Opposite Gender.

Dependent Variable	Decision of Noncompliance in Given Month					
	Siblings		Sibs in Other Schools		Sibs of Opposite Sex	
	Own Max	Sib Max	Own Max	Sib Max	Own Max	Sib Max
	(1)	(2)	(3)	(4)	(5)	(6)
Frac of Own Peers Warned	-0.0208*	-0.0018	-0.0207**	-0.0119	-0.0196**	0.0054
	(0.012)	(0.022)	(0.009)	(0.017)	(0.010)	(0.024)
Frac of Own Peers Warned (Lag)	-0.0031	0.0145	-0.005	-0.0079	-0.009	-0.0151
	(0.013)	(0.024)	(0.010)	(0.019)	(0.010)	(0.026)
Max of Sibs Peers Warned	-0.0037	-0.0376***				
	(0.024)	(0.011)				
Max of Sibs Peers Warned (Lag)	-0.0239	-0.0413***				
	(0.025)	(0.012)				
Max of Sibs Peers Warned			0.0028	-0.0589***		
(In Other Schools)			(0.025)	(0.010)		
Max of Sibs Peers Warned (Lag)			-0.0307	-0.0507***		
(In Other Schools)			(0.026)	(0.011)		
Max of Sibs Peers Warned					0.0057	-0.0374***
(Sib of Opposite Sex)					(0.020)	(0.013)
Max of Sibs Peers Warned (Lag)					0.0046	-0.0179
(Sib of Opposite Sex)					(0.023)	(0.015)
Warnings Stage 1	-0.0157***	-0.0151***	-0.0161***	-0.0148***	-0.0170***	-0.0140***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Warnings Stage 2	-0.0425***	-0.0377***	-0.0415***	-0.0375***	-0.0435***	-0.0355***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Warnings Stage 3	-0.0528***	-0.0542***	-0.0562***	-0.0542***	-0.0550***	-0.0537***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Warnings Stage 4	-0.1335***	-0.1094***	-0.1243***	-0.1088***	-0.1252***	-0.1042***
	(0.020)	(0.017)	(0.019)	(0.017)	(0.018)	(0.019)
Warnings Stage 5	-0.1397*	0.0691	-0.0547	0.0688	-0.1065*	0.0983
	(0.074)	(0.081)	(0.081)	(0.080)	(0.055)	(0.106)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
HH-FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1172700	1176678	1172700	1176678	1172700	1176678
R-squared	0.13	0.13	0.12	0.13	0.12	0.13

Notes: Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01. Included controls are age dummies, birth order dummies, month and year dummies and number of brothers and sisters in the age categories 0 to 5, 6 to 10, 11 to 15, and 16 to 18. All specifications include family fixed effects. The number of observations is reduced as we only include families with at least two children.

Table 6: Effect of Siblings' Peers Receiving a Justification for their Failure on a Child's Attendance Behavior (Failure to Attend a Sufficient Number of Days in a Month).

Dependent Variable	**Decision of Noncompliance in Given Month**			
	Siblings		Siblings in Other Schools	
	Own Max	Sib Max	Own Max	Sib Max
	(1)	(2)	(3)	(4)
Warning 1	-0.0270***	-0.0252***	-0.0241***	-0.0242***
	(0.002)	(0.002)	(0.002)	(0.002)
Warning 2	-0.0628***	-0.0584***	-0.0560***	-0.0550***
	(0.005)	(0.005)	(0.004)	(0.004)
Warning 3	-0.0689***	-0.0634***	-0.0639***	-0.0590***
	(0.01)	(0.01)	(0.009)	(0.009)
Warning 4	-0.1231***	-0.1213***	-0.1123***	-0.1107***
	(0.019)	(0.019)	(0.017)	(0.017)
Warning 5	-0.0499	-0.0586	-0.0543	-0.054
	(0.08)	(0.071)	(0.078)	(0.077)
Frac of Own Peers Justif (lag 1)	0.0251	0.1129***	0.0363**	0.0401**
	(0.019)	(0.024)	(0.018)	(0.018)
Frac of Own Peers Justif (lag 2)	0.0591**	0.0491*	0.0623**	0.0488*
	(0.026)	(0.028)	(0.026)	(0.026)
Max Frac of Sib Peers Justified (lag 1)	-0.0309	-0.0204		
	(0.019)	(0.018)		
Max Frac of Sib Peers Justified (lag 2)	-0.0465**	-0.0379*		
	(0.023)	(0.021)		
Max Frac of Sib Peers in Other Schools (lag 1)			-0.0466*	-0.0325
			(0.024)	(0.021)
Max Frac of Sib Peers in Other Schools (lag 2)			-0.0910***	-0.0648***
			(0.024)	(0.023)
Controls	Yes	Yes	Yes	Yes
HH-FE	Yes	Yes	Yes	Yes
Observations	1190067	1190608	1207334	1212400
R-squared	0.15	0.16	0.15	0.15

Notes: Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01. Included controls are age dummies, birth order dummies, month and year dummies and number of brothers and sisters in the age categories 0 to 5, 6 to 10, 11 to 15, and 16 to 18. All specifications include family fixed effects. The number of observations is reduced as we only include families with at least two children.

Table 7: Effect of Own Peers Receiving a Justification for their Failure on a Child's Attendance Behavior (Failure to Attend a Sufficient Number of Days in a Month).

Dependent Variable	**Decision of Noncompliance in Given Month**				
	(1)	(2)	(3)	(4)	(5)
Frac of Own Peers Justified	-0.0139				
	(0.016)				
Lag of Frac of Own Peers Justified	0.0444***	0.0131			
	(0.016)	(0.016)			
Lag 2 of Frac of Own Peers Justified	0.0507**	0.0117			
	(0.024)	(0.022)			
Any Own Peer Justified			0.0074**	0.0061**	0.0048
			(0.003)	(0.003)	(0.003)
Lag of Any Own Peer Justified				0.0057**	0.0049*
				(0.003)	(0.003)
Frac of Peers Fail		1.1888***	1.1902***	1.1908***	1.1386***
		(0.021)	(0.021)	(0.021)	(0.021)
Lag of Frac of Peers Fail					0.1136***
					(0.014)
Warning 1	-0.0234***	-0.0209***	-0.0185***	-0.0184***	-0.0208***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Warning 2	-0.0539***	-0.0486***	-0.0437***	-0.0436***	-0.0484***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Warning 3	-0.0597***	-0.0597***	-0.0528***	-0.0527***	-0.0587***
	(0.009)	(0.008)	(0.007)	(0.007)	(0.008)
Warning 4	-0.1144***	-0.0955***	-0.1018***	-0.0965***	-0.1020***
	(0.017)	(0.014)	(0.013)	(0.013)	(0.014)
Warning 5	-0.0493	-0.09	-0.1101*	-0.0789	-0.0992
	(0.073)	(0.067)	(0.063)	(0.065)	(0.07)
Controls	Yes	Yes	Yes	Yes	Yes
HH-FE	Yes	Yes	Yes	Yes	Yes
Observations	1227526	1227526	1245129	1243597	1215057
R-squared	0.13	0.23	0.23	0.23	0.24

Notes: Robust standard errors in parentheses. * p<0.1 ** p<0.05 *** p<0.01. Included controls are age dummies, birth order dummies, month and year dummies and number of brothers and sisters in the age categories 0 to 5, 6 to 10, 11 to 15, and 16 to 18. All specifications include family fixed effects.